

Voorbeeld modules voor Datamanagement paragraaf – vs. 2.0

Steeds meer subsidiegevers vragen om een korte paragraaf over datamanagement op te nemen in de subsidieaanvraag. Om onderzoekers te ondersteunen bij dit schrijfproces hebben wij hieronder een aantal voorbeeld modules geplaatst die gebruikt kunnen worden in de datamanagement paragraaf. Daarbij benadrukken we wel dat ieder onderzoek zijn eigen unieke onderzoeksgegevens heeft en dat de onderstaande modules slechts een startpunt zijn voor de datamanagement paragraaf. Werk dit verder uit, waar nodig. De modules zijn gestructureerd naar de vier vragen over datamanagement die o.a. door de subsidiegever NWO worden gesteld. Ook als er door jouw subsidiegever wordt gevraagd om een algemene datamanagement paragraaf, probeer dan wel alle antwoorden op deze vier vragen aan bod te laten komen.

In vierkante haken een voorbeeldtekst zoals je deze kan opnemen in een datamanagement paragraaf. Schuingedrukt eventuele opties die al dan niet van toepassing zijn voor jouw onderzoek.

Julika Verhoeven (CvVO)
Ivo Schrijers (CvVO)
Liesbeth Hoogkamp (ICT)
Imke Adams (SUOO)

1. Worden er data verzameld of gegenereerd die voor hergebruik geschikt zijn?

Ja: beantwoord dan vraag 2 t/m 4.

Nee: licht dan toe waarom het onderzoek niet leidt tot herbruikbare data, of tot data die niet kunnen worden opgeslagen of om andere redenen niet relevant zijn voor hergebruik.

[In dit project wordt gebruik gemaakt van *nieuw te verzamelen data/ bestaande databases* .
De verzamelde data is *wel/niet* geschikt voor hergebruik.]

Indien de data niet geschikt is voor hergebruik, overweeg goed waarom jouw data niet herbruikbaar zou zijn. Data kan geanonimiseerd worden en als dat niet mogelijk is kunnen eventueel groepsgemiddelden (bijv. stratificatie) wel gepubliceerd worden.

2. Waar worden de data gedurende het onderzoek opgeslagen?

[De HAN heeft een Datamanagementplan opgesteld wat van toepassing is (*zie bijlage ...*). Gedurende het onderzoek worden ruwe en bewerkte data opgeslagen op de speciaal voor onderzoeksdata gecreëerde veilige netwerkschijf (R-schijf) van HAN. Deze fungeert als ‘kluis’: tot de projectomgeving op deze schijf hebben gedurende het project slechts de door de projectleider toegewezen onderzoekers toegang, alsmede de lector en het lectoraatssecretariaat. De R-schijf heeft een vaste indeling, met een mappenstructuur gebaseerd op de fasen van een lopend onderzoek. Er is een aparte map beschikbaar voor de sleutels van versleutelde data. De mappenstructuur ziet er als volgt uit:

Projectnummer-projectnaam

- A. Ruwe datasets
- B. Halfbewerkte datasets
- C. Datasets voor analyse
- D. Publicaties
- E. Contact en Contract

Projectnummer -sleutelgegevens

- A. Sleutelgegevens (map met aparte vergrendeling en zeer beperkte toegang)]

[Fysieke data (zoals labjournaals, ingevulde papieren vragenlijsten, weefsels, kweken) worden opgeslagen in ... en ...jaar bewaard]. (Beschrijf ook hoe je dataverlies voorkomt)

[De resultaten van andere partners van dit project worden bij de betreffende partners opgeslagen, op de volgende wijze]

3. Hoe worden de data na afloop van het project voor de lange termijn opgeslagen en voor hergebruik beschikbaar gesteld voor derden? Voor wie zijn de data toegankelijk?

[Alle datasets worden gedurende 10 jaar bewaard na het eind van het onderzoek op de R-schijf, indien nodig versleuteld. Alleen een vooraf aangewezen supervisor en vervanger hebben toegang tot de mappen van een lectoraat.]

[De publicaties en ... (*eventuele andere producten*) worden zoveel mogelijk open access gepubliceerd.]

(Denk dan ook aan het opnemen van budget, voor het open acces publiceren van je onderzoeksresultaten, in de subsidieaanvraag).

Indien verzamelde data wel beschikbaar wordt gesteld voor hergebruik:

[Om data beschikbaar te stellen voor derden zullen geanonimiseerde data beschikbaar worden gesteld via Er wordt gebruikt gemaakt van een digital object identifier (doi) om duurzaam naar data te verwijzen]*.

[In de samenwerkingsovereenkomst van de HAN en zijn afspraken over databeheer en (her)gebruik vastgelegd. Daarin zijn de volgende afspraken opgenomen ...]

Indien verzamelde data niet beschikbaar wordt gesteld voor hergebruik:

[In overleg met de partners is besloten de ruwe data niet als open source beschikbaar te maken gezien de relatief makkelijke herleidbaarheid van data naar personen.]

(overweeg hierbij of publicatie van bijv. groepsgemiddelden door middel van stratificatie wel mogelijk is).

* (De HAN stimuleert het (deels) open access publiceren van datasets. Dit kan via DANS of via een andere open repository met het data seal of approval.)

4. Welke voorzieningen (ICT, (beveiligd) archief, koelkasten of juridische expertise) zijn naar verwachting nodig voor de opslag van data gedurende het onderzoek en na het onderzoek? Zijn deze beschikbaar?* * Met ICT-voorzieningen voor dataopslag wordt bedoeld zaken als capaciteit voor dataopslag, bandbreedte voor datatransport en rekenkracht voor dataverwerking.

[ICT-voorzieningen zijn centraal geregeld door het HAN Servicebedrijf voor de hele HAN. Gedurende het onderzoek worden ruwe en bewerkte data opgeslagen op de speciaal voor onderzoeksdata gecreëerde veilige netwerkschijf (R-schijf) van HAN. Alle datasets worden gedurende 10 jaar bewaard na het eind van het onderzoek op de R-schijf, indien nodig versleuteld. Binnen dit project zijn geen additionele voorzieningen nodig.]

(Geef hier ook aan of er gebruik wordt gemaakt van voorzieningen van partners).

Nuttige bronnen:

<http://rdm.uva.nl/plannen/eisen-van-financiers/eisen-van-financiers.html>

<https://www.surf.nl/diensten-en-producten/data-archive/index.html>

<https://dans.knaw.nl/nl>